## Caenorhabditis elegans embryonic axial patterning requires two recently discovered posterior-group Hox genes

Kimberly Van Auken, Daniel C. Weaver\*, Lois G. Edgar, and William B. Wood<sup>†</sup>

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309-0347

Contributed by William B. Wood, February 7, 2000

Hox genes encode highly conserved transcription factors that control regional identities of cells and tissues along the developing anterior-posterior axis, probably in all bilaterian metazoans. However, in invertebrate embryos other than Drosophila, Hox gene functions remain largely unknown except by inference from sequence similarities and expression patterns. Recent genomic sequencing has shown that Caenorhabditis elegans has three Hox genes of the posterior paralog group [Ruvkun, G. & Hobert, O. (1998) Science 282, 2033-2041]. However, only one has been previously identified genetically, and it is not required for embryonic development [Chisholm, A. (1991) Development (Cambridge, U.K.) 111, 921-932]. Herein, we report identification of the remaining two posterior paralogs as the nob-1 gene and the neighboring php-3 gene. Elimination of nob-1 and php-3 functions causes gross embryonic defects in both posterior patterning and morphogenetic movements of the posterior hypodermis, as well as posterior-toanterior cell fate transformations and lethality. The only other Hox gene essential for embryogenesis is the labial/Hox1 homolog ceh-13, required for more anterior patterning [Brunschwig, K., Wittmann, C., Schnabel, R., Burglin, T. R., Tobler, H. & Muller, F. (1999) Development (Cambridge, U.K.) 126, 1537-1546]. Therefore, essential embryonic patterning in C. elegans requires only Hox genes of the anterior and posterior paralog groups, raising interesting questions about evolution of the medial-group genes.

Recent genomic comparisons of several invertebrates including *Caenorhabditis elegans* support the view that expansion of the Hox genes to include most or all of the present paralog groups occurred early in bilaterian evolution before the divergence of protostomes and deuterostomes (1). Previous work in several laboratories identified a small loosely linked cluster of four *C. elegans* Hox genes (Fig. 1a): the anterior-group homolog *ceh-13*, two medial-group homologs of the *Antp* class, *lin-39* and *mab-5*, and the posterior-group homolog *egl-5* (2, 3). The expression patterns and postembryonic functions of these genes are consistent with roles in postembryonic specification of regional identities (4, 5). However, only *ceh-13* is essential for embryonic development (6), whereas *egl-5* is not (7), supporting an earlier suggestion that additional posterior-group genes might remain to be discovered (3).

To search genetically for such genes, we have performed screens for embryonic and larval lethal mutations that result in severe posterior morphological abnormalities without disrupting general aspects of development such as cell proliferation and tissue specification. We have found mutations representing five complementation groups that result in such abnormalities, which we have designated the "No back end" or Nob phenotype (ref. 8; L.G.E., S. Carr, H. Wang, and W.B.W., unpublished work). We report herein on nob mutations affecting two previously uncharacterized Hox genes, recently also identified by genomic sequencing (9), which belong to the posterior paralog group and seem to be essential for embryonic development and posterior patterning.

## Materials and Methods

Cloning and Characterization of nob-1 and php-3. C. elegans strains were handled by standard methods (10). The nob-1(ct223) mutation was mapped to a locus between spe-6 and pie-1 on the right arm of chromosome III (LGIIIR) by standard three-factor mapping (11). The ct230 and ct351 mutations were identified in screens for posterior patterning genes on LGIII after mutagenesis with trimethylpsoralen (8); complementation tests with ct223 defined these as alleles of nob-1. Representational difference analysis was performed as described (12) with one modification: given the relative simplicity of the C. elegans genome, genomic DNAs were digested with a four-cutter restriction enzyme, Sau3AI, resulting in average fragments of approximately 300 bp. Difference products from the representational difference analysis experiments were cloned into pBluescript SK(+) (Stratagene) and sequenced on an Applied Biosystems sequencer. The nob-1(ct230) lesion was identified by direct sequencing of PCR products generated by a single-worm PCR assay (13) with primers recognizing sequences in exon 1 and intron 2 of nob-1. The endpoints of the ct223 and ct351 deletions were also determined by single-worm PCR with primers recognizing genomic sequence in the regions flanking nob-1 and php-3. nob-1 cDNA clones yk467d4.5 (nob-1A; GenBank accession no. C49108) and yk403d9.5 (nob-1B; C45910) were sequenced with T3 and T7 primers. The php-3 transcript was detected by reverse transcription-PCR with mixed-stage RNA prepared by the Trizol (GIBCO) method. PCR primers used contained EcoRI restriction sites such that the products could be subcloned into pBluescript SK(+) for sequencing. The 5' ends of nob-1 and php-3 transcripts were determined by rapid amplification of cDNA ends with the 5' Rapid Amplification of cDNA Ends Anchor Primer (GIBCO); both transcripts were found to carry the SL1 splice leader sequence (12). Comparison of the complete cDNA and genomic sequences (GenBank) showed that the structure of nob-1 is different from that predicted originally by GENEFINDER (15); we have submitted the cDNA sequences to GenBank with the accession numbers: AF172090 for nob-1A, AF172091 for nob-1B, and AF172092 for php-3.

**RNA-Mediated Gene Interference (RNAi).** RNAi experiments were performed as described (16). cDNA clones from *nob-1* and *php-3* were amplified with T3 and T7 primers and purified by phenol/chloroform extraction. RNA was synthesized with T3 and T7 RNA polymerases (GIBCO/BRL) and injected into wild-type or *nob-1(ct230)* hermaphrodites at concentrations of 0.5–1.0 mg/ml. To determine RNAi effects on male development, RNAs

Abbreviation: kb, kilobase; GFP, green fluorescent protein; RNAi, RNA-mediated interference.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF172090 for nob-1A, AF172091 for nob-1B, and AF172092 for php-3).

<sup>\*</sup>Present address: Genomica, Inc., Boulder, CO 80303-7817

<sup>&</sup>lt;sup>†</sup>To whom reprint requests should be addressed. E-mail: wood@stripe.colorado.edu.

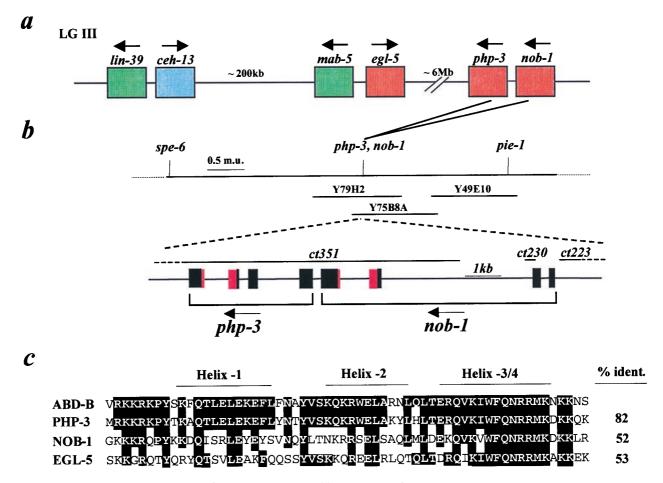


Fig. 1. Mapping, cloning, and sequence analysis of *nob-1* and *php-3* genes. (a) Arrangement of the *C. elegans* Hox genes. The anterior-group gene is indicated in blue; medial-group genes are indicated in green; and posterior-group genes are indicated in red. Arrows indicate directions of transcription. kb, kilobase; Mb, megabase. (b) Genomic organization and mutant lesions of *nob-1* and *php-3*. Partial genetic and physical maps show location of *nob-1* and *php-3* on LGIIIR and on yeast artificial chromosomes Y79H2 and Y75B8A. The extent of each deletion mutation is shown above the genes. Dots extending toward the right and left under *ct223* and *ct351*, respectively, indicate that these deletions extend beyond the boundaries of the figure (see text). m.u., map units (centimorgans). (c) Sequence alignment of predicted homeodomains encoded by *Drosophila AbdB* and *C. elegans* posterior-group genes *php-3*, *nob-1*, and *egl-5*. Shading indicates identities between *AbdB* and the *C. elegans* genes is shown at the right.

were injected into him-8(e1489) hermaphrodites, which produce about 40% male self progeny.

**Phenotypic Analysis.** The pal-1::lacZ and ceh-13::GFP (GFP = green fluorescent protein) reporters were maintained in a strain of genotype nob-1(ct223)/nob-1(ct223);eDp6, where eDp6 is a free duplication that provides a wild-type copy of nob-1 and php-3 and is lost meiotically at a frequency of  $\approx 50\%$ . For assaying  $\beta$ -galactosidase activity, embryos were fixed and stained as described (17). GFP expression was observed with a Leitz microscope equipped with Nomarski optics and epifluorescence. Lineage analysis was performed with a multifocal-plane time-lapse video recording system (18).

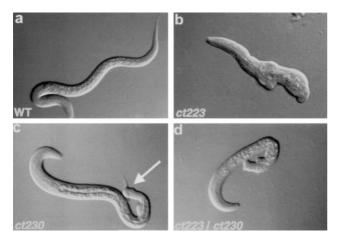
## Results

**Genetic Characterization and Cloning of the** *nob-1* **Gene.** Three recessive mutations identified in Nob screens defined a complementation group given the gene name *nob-1*. Two of these, *nob-1(ct223)* and *ct351*, seem phenotypically identical, resulting in 100% larval lethality and little or no posterior morphogenesis (Fig. 2). The third, weaker allele, *ct230*, results in 3% Nobs and 97% viable animals with variably misshapen tails (Vab pheno-

type). Heteroallelic *ct230/ct223* or *ct230/ct351* animals have intermediate morphogenesis and viability.

We mapped *nob-1* genetically to a gap in the physical map that until recently contained no ordered cosmid or yeast artificial chromosome clones; thus, standard approaches to positional cloning were impractical. Instead, we used the subtractive technique of representational difference analysis (12), suspecting that the trimethylpsoralen-induced allele ct230 might be a small deletion (8). We were able to identify a 450-bp Sau3AI fragment present in wild-type (N2) but not in ct230 DNA. Its sequence matched the 5' ends of two cDNA clones in the C. elegans Expressed Sequence Tag database, yk467d4.5 (nob-1A) and yk403d9.5 (nob-1B). Sequencing of these clones and comparison to genomic sequence indicated that they represented transcripts of a Hox gene (Fig. 1b) with strong similarity in the predicted homeodomain to Drosophila AbdB and other members of the posterior paralog group (Fig. 1c). We confirmed that these transcripts derive from nob-1 by sequencing genomic DNA from nob-1(ct230) and showing that it contains a 192-bp deletion that removes part of the second exon and intron (Fig. 1b). Nob-1A is a 1.35-kb transcript that encodes a complete homeodomain in exons 3 and 4, and Nob-1B is an alternatively spliced 2.2-kb

4500 | www.pnas.org Van Auken *et al.* 



**Fig. 2.** Wild-type and defective *nob-1* first-stage (L1) larvae. Nomarski images are shown with anterior to the left. (a) Wild type. (b) Homozygous *ct223* as well as *ct351* larvae have the somewhat variable Nob phenotype and terminally arrest development at this stage. (c) Homozygous *ct230* larvae show a less severe phenotype, with variably misshapen tails (arrow), and survive to become fertile adults. (d) Heteroallelic animals have intermediate defects and viability.

transcript that includes an extended exon 2 as well as intron 3; its predicted translation product lacks the homeodomain.

nob-1 Is Flanked by Another Hox Gene, php-3. Identification of nob-1 sequence allowed the *C. elegans* Genome Sequencing Consortium to position nob-1 physically on a group of previously unmapped yeast artificial chromosome clones and to close the gap in the physical map (A. Coulson, personal communication). Their further analysis (15) completed the genomic sequence of nob-1 (designated Y75B8A.2) and identified a second posterior-group Hox gene (Y75B8A.1), beginning just 232 bp from the 3' end of nob-1 (Fig. 1b) and with considerably higher similarity than nob-1 to AbdB in the predicted homeodomain (ref. 1; Fig. 1c). We named this gene php-3 (posterior Hox paralog 3; 1 and 2 are egl-5 and nob-1, respectively). php-3 and nob-1 have similar gene structures (Fig. 1b) and have 47% nucleotide identity throughout their coding regions.

Functions of nob-1 and php-3 in Embryonic Patterning. To investigate the basis for the lethality resulting from ct223 and ct351, we identified the molecular lesions in each (Fig. 1b). ct351 is a large deletion that removes part of *nob-1*, including the homeobox, and all of php-3, in addition to at least 15 kb downstream of php-3. ct223 is a nonoverlapping deletion of  $\approx$ 35 kb that lies entirely upstream of nob-1 coding sequence and deletes presumed regulatory sequences. Additional predicted genes are also removed by these deletions: two upstream by ct351 and two downstream by ct223. However, because the resulting phenotypes are apparently identical, it seems likely that both deletions cause loss of both nob-1 and php-3 activities and that the observed defects result from the loss of these gene functions. We have no evidence that these genes constitute an operon (19) (both mRNAs carry the 5' splice-leader sequence SL1), but their separation by only 232 bp suggests that regulatory elements for the downstream gene *php-3* could be present upstream of *nob-1*.

In contrast to the moderately abundant *nob-1* transcripts, we could identify transcripts of *php-3* only by reverse transcription–PCR and not by RNA blotting or searches of the Expressed Sequence Tag database, suggesting that they are rare. In genetic experiments, *nob-1(ct230)* behaves as a hypomorph rather than a null allele, and reverse transcription–PCR analysis indicated that it does not eliminate *php-3* transcription (data not shown).

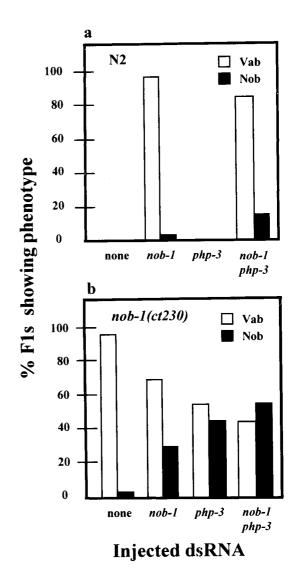


Fig. 3. RNAi phenotypes. Percentages of  $F_1$  animals with a Nob or Vab phenotype are shown for wild-type (a) or nob-1(ct230) (b) hermaphrodites after injection with double-stranded (ds)RNA sequences from nob-1, php-3, or both. In similar experiments with a him-8 (high-incidence-of-males) strain in place of N2, php-3(RNAi) alone also had no apparent effect on male tail development. Progeny ( $n = \infty 500$ ) were scored 24–96 h after each injection. Negative controls ("none") are uninjected animals. It is very unlikely that php-3(RNAi) could be crossinterfering with nob-1 function or vice versa, because this phenomenon has been observed only when there is >80% identity in overall nucleotide sequence (30).

Because ct223 and ct351, which behave genetically as null alleles, are likely to result in loss of both gene functions, we used RNAi (20) in attempts to define separate loss-of-function (lf) phenotypes for nob-1 and php-3. As shown in Fig. 3a, injection of nob-1 RNA [nob-1(RNAi)] into wild-type hermaphrodites resulted in 3% Nobs and 97% Vabs, much like the nob-1(ct230) phenotype. php-3(RNAi) alone seemed to have no effect, but when combined with nob-1(RNAi), it increased the proportion of Nobs to 15%. Injection of these RNAs into homozygous nob-1(ct230) animals (Fig. 3b) showed more striking effects; in particular, php-3(RNAi) alone resulted in almost 50% of progeny with the presumed null (Nob) phenotype. Similar effects were observed on development of males after injections into him-8 (high incidence of males) hermaphrodites (data not shown). Thus, both nob-1(RNAi) and php-3(RNAi) can enhance the ct230

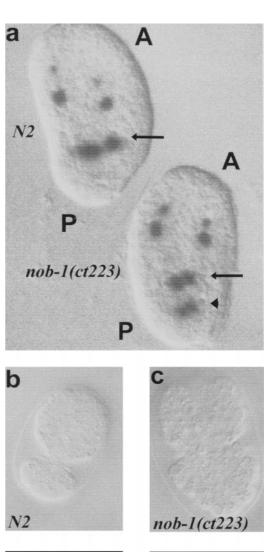
defective phenotype, confirming that ct230 is a hypomorph and indicating that loss of php-3 function can affect posterior patterning. Quantitative interpretation of these results is problematic, because silencing by RNAi can be difficult for genes whose transcripts are rare, like php-3; moreover, we have preliminary evidence (not shown) that nob-1 is expressed predominantly in the nervous system, one of the tissues in which RNAi has proven least effective (20, 21). These factors could account for our observations that php-3(RNAi) alone apparently does not affect embryogenesis and that the combination of nob-1(RNAi) and php-3(RNAi) causes a less severe phenotype than either the ct223 or ct351 deletion. More rigorous definition of individual nob-1 and php-3 functions must await analysis of additional deletion mutations. The above results suggest that these genes work together in posterior embryonic patterning and that php-3 may play a relatively minor role.

To obtain preliminary insight into their combined functions, we have begun to analyze the defective phenotype of nob-1(ct223). A hallmark of loss-of-function mutations in Hox genes is the resulting posterior-to-anterior transformations in regional cell and tissue fates (22). To determine whether mutations affecting *nob-1* result in such transformations during embryonic development, we examined ct223 embryos for expression of two region-specific markers: (i) a lacZ reporter for the caudal homolog pal-1, also required for embryonic posterior development in C. elegans (ref. 23; L.G.E., S. Carr, H. Wang, and W.B.W., unpublished work), and (ii) a GFP reporter for *ceh-13*. The results suggest posterior-to-anterior transformations in both gut and nervous system. The pal-1 reporter, normally expressed in only two gut cells in the central region during mid embryogenesis, is seen in two additional posterior gut cells of mutant embryos (Fig. 4a). The ceh-13 reporter, normally expressed in a few mid-ventral neuronal precursors (ref. 6; Fig. 4 b and d) is seen in at least six additional neurons from more posterior lineages of mutant embryos (Fig. 4 c and e).

We also examined the posterior embryonic lineages for all major tissues (hypodermis, muscle, gut, and neurons) in embryos homozygous for the ct223 mutation (n = 4) to determine whether cell fate changes might be reflected in altered lineage patterns. We found that almost all aspects of the cell lineage (which is completed during the first half of embryogenesis before hypodermal enclosure; ref. 24), including the number, timing, and orientation of cell divisions (to be described elsewhere), are normal. The exception is in the gut, where there is a minor posterior-to-anterior transformation in the lineal pattern, consistent with the expression results described above (descendants of Eprp divide with the pattern of Earp, which is normally different). Analysis of the cell migrations and shape changes required for morphogenesis (25) identified a failure of posterior hypodermal cells to enclose and elongate the embryo properly, suggesting that these cell fates are also improperly specified. Thus, nob-1 and php-3 are not generally required for dictating the pattern of cell divisions during embryogenesis, but like their homologs in other embryos, they are needed for execution of appropriate axial fates.

## **Discussion**

It is now known from the essentially complete genomic sequence (15) that *C. elegans* contains a total of six Hox genes (Fig. 1a). Of these, only *ceh-13* and the *nob-1/php-3* pair are essential for viability and normal embryonic morphogenesis. In agreement with this conclusion, we have observed no differences in the phenotypes of *nob-1(ct223)* single mutants and *mab-5(lf);nob-1(ct223)* and *egl-5(lf);nob-1(ct223)* double mutants (data not shown). Moreover, animals triply mutant for *lin-39, mab-5*, and *egl-5* can be viable (ref. 26; A. Chisholm, personal communication) ruling out required but redundant functions in embryogenesis. Thus, the patterning essential for *C. elegans* embryo-



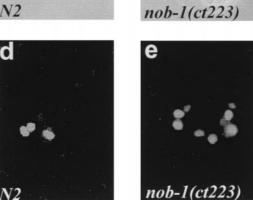


Fig. 4. Posterior-to-anterior fate transformations in nob-1(ct223) embryos. (a-c) Nomarski images with anterior (A) toward the top and posterior (P) toward the bottom. (d and e) Epifluorescence images of b and c, respectively. In a, pal-1::lacZ expression is shown in a wild-type (upper) and a nob-1(ct223) (lower) embryo. The arrows point to the two mid-region gut cells that express the reporter in both wild-type and mutant embryos; the arrowhead indicates the ectopic expression in two more posterior gut cells of the nob-1(ct223) embryo. Analysis of several such embryos with a pal-1::GFP as well as the pal-1::lacZ reporter identified the ectopically expressing cells as those of the int8/9 intestinal segment. The four additional anterior expressing cells are myoblasts not affected in nob-1(ct223). (b and c) Ventral views of wild-type and nob-1(ct223) embryos, respectively, during early morphogenesis when the embryo is enclosing. (d and e) Expression of the ceh-13::GFP reporter in mid-ventral neurons of the same two embryos, respectively. Lineaging shows that the ectopically expressing cells in such Nob embryos are derived from more posterior lineages than the normally expressing cells. (This anteriorization is not evident from the figure, because at this stage, the Nob embryo is already morphologically abnormal.)

4502 | www.pnas.org Van Auken *et al.* 

genesis requires at most three Hox genes, representing only the anterior (labial/Hox1) and posterior (AbdB/Hox9-13) paralog groups. In this regard, C. elegans is different from both Drosophila, another protostomic ecdysozoan (27), and vertebrates on the deuterostome branch of the bilaterian tree, which also require medial-group genes (22). The major Hox paralog classes must have arisen before protostome-deuterostome divergence (1, 28). Therefore, we have no basis for deciding whether the C. elegans mechanism is primitive or derived, that is, whether the common ancestor required only anterior and posterior group genes and only later co-opted medial group genes for embryonic roles, or whether it required medial group genes as well, which later became unnecessary for embryogenesis in the nematode line of descent. In Cnidarians, which occupy a basal position in animal phylogeny, only anterior-group and posterior-group Hox genes have been identified (29). Our findings therefore raise intriguing questions about the evolution of the medial group

- de Rosa, R., Grenier, J. K., Andreeva, T., Cook, C. E., Adoutte, A., Akam, M., Carroll, S. B. & Balavoine, G. (1999) Nature (London) 399, 772–776.
- 2. Kenyon, C. & Wang, B. (1991) Science 253, 516-517.
- Burglin, T., Ruvkun, G., Coulson, A., Hawkins, N., McGhee, J., Schaller, D., Wittmann, C., Muller, F. & Waterston, R. (1991) Nature (London) 351, 703 (lett.).
- Wang, B., Muller-Immergluck, M., Austin, J., Robinson, N., Chisholm, A. & Kenyon, C. (1993) Cell 74, 29–42.
- 5. Clark, S., Chisholm, A. & Horvitz, H. (1993) Cell 74, 43-55.
- Brunschwig, K., Wittmann, C., Schnabel, R., Burglin, T. R., Tobler, H. & Muller, F. (1999) Development (Cambridge, U.K.) 126, 1537–1546.
- 7. Chisholm, A. (1991) Development (Cambridge, U.K.) 111, 921–932.
- Yandell, M., Edgar, L. & Wood, W. (1994) Proc. Natl. Acad. Sci. USA 91, 1381–1385.
- 9. Ruvkun, G. & Hobert, O. (1998) Science 282, 2033-2041.
- 10. Brenner, S. (1974) Genetics 77, 71-94.
- Sulston, J. & Hodgkin, J. (1988) in *The Nematode Caenorhabditis elegans*, ed. Wood, W. B. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 81–122.
- 12. Lisitsyn, N., Lisitsyn, N. & Wigler, M. (1993) Science 259, 946-951.
- Williams, B., Schrank, B., Huynh, C., Shownkeen, R. & Waterston, R. (1992) Genetics 131, 609–624.
- 14. Krause, M. & Hirsh, D. (1987) Cell 49, 753-761.
- 15. *C. elegans* Sequencing Consortium (1998) *Science* **282**, 2012–2018.
- Montgomery, M. K., Xu, S. & Fire, A. (1998) Proc. Natl. Acad. Sci. USA 95, 15502–15507.

genes and about whether other bilaterian invertebrates have the *C. elegans* mode of Hox gene function in embryogenesis. Meanwhile, our results indicate the importance of determining function experimentally rather than assuming it from sequence similarity. From sequence analysis (1), the *AbdB* ortholog in *C. elegans* seems to be *php-3*, which our preliminary findings suggest may have only a minor role in posterior patterning compared with the role of the more divergent *nob-1*.

We are grateful to Y. Kohara for expressed sequence tag sequences and cDNA clones, to F. Mueller for the *ceh-13::GFP* construct, to A. Coulson for helpful discussions regarding sequencing of the *nob-1* region, to the *C. elegans* Sequencing Consortium for genomic sequence information, to C. Cary, F. Notka, and B. Robertson for assistance with some experiments, and to B. Parr, M. Scott, W. Li, and other members of our laboratory for comments on the manuscript. This research was supported by National Institutes of Health Grant HD-14958 to W.B.W.

- Edgar, L. G. (1995) in Caenorhabditis elegans: Modern Biological Analysis of an Organism, eds. Epstein, H. F. & Shakes, D. C. (Academic, New York), pp. 303–322.
- 18. Thomas, C., DeVries, P., Hardin, J. & White, J. (1996) Science 273, 603-607.
- Zorio, D., Cheng, N., Blumenthal, T. & Spieth, J. (1994) Nature (London) 372, 270–273
- Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. (1998) *Nature (London)* 391, 806–811.
- 21. Montgomery, M. K. & Fire, A. (1998) Trends Genet. 14, 255-258.
- Manak, J. & Scott, M. (1994) Development (Cambridge, U.K.) 120, Suppl., 61–71
- 23. Hunter, C. & Kenyon, C. (1996) Cell 87, 217–226.
- Sulston, J., Schierenberg, E., White, J. & Thomson, J. (1983) Dev. Biol. 100, 64-119.
- 25. Priess, J. R. & Hirsh, D. I. (1986) Dev. Biol. 117, 156-173.
- Wrischnik, L. A. & Kenyon, C. J. (1997) Development (Cambridge, U.K.) 124, 2875–2888
- Aguinaldo, A. M. A., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey,
  J. R., Raff, R. A. & Lake, J. A. (1997) *Nature (London)* 387, 489–493.
- 28. Knoll, A. H. & Carroll, S. B. (1999) Science 284, 2129-2137.
- Martinez, D. E., Bridge, D., Masuda-Nakagawa, L. M. & Cartwright, P. (1998) Nature (London) 393, 748–749.
- 30. Tabara, H., Grishok, A. & Mello, C. C. (1998) Science 282, 430-431.